Quality in measurement: beyond the deployment barrier

Tony McGregor
The University of Waikato and NLANR
Computer Science Department
Private Bag 3105
Hamilton
New Zealand
tonym@nlanr.net*

Abstract

Network measurment stands at an intersection in the development of the science. We explore possbile futures for the area and propose some guidelines for the development of stronger measurement techniques. The paper concludes with a discussion of the work of the NLANR and WAND network measurement groups including the NLANR Network Analysis Infrastructure, AMP, PMA, analysis of Voice over IP traffic and seperation of HTTP delays into queuing delay, network latency and server delay.

1 Introduction

In 1896, a new science was born that would, arguably, have more impact on mankind than any other before it. In that year, Henri Becquerel, having been intrigued by the reacent discovery of x-rays, studied the phosphoresence of uranium salts. Instead of finding the x-rays he expected he discovered what Marie and Pierre Curie two years later called radioactivity. In 1900 Ernest Rutherford and Frederick Soddy discovered that elements that are radioactive changed into other elements. Over the next 30 years a new science was established; *modern physics* which lead to the discovery of the strcture of the attom, quantum mechanics, demonstration of relitivity and, of course, the world of nuclear wepons, power and medicine.

In its early years moden physics was pursuied entyerley for its own bennifit. Initially there was very little sign of any practical outcome. However in October 1939, an external event (WWII) and the vision of a group of scientists that included Eienstein inspired a Roosevelt to invest in the new science because of its potential use in weponary. In just 49 years Becquerel's accidental discovery of raidoactivity lead to "Little Boy" being dropped on Hiroshima and the beginnig of the nuclear age and all that that has entaied.

Over those 50 years the science of nuclear physics evolved from an interesting research problem to main stream science. A few scientests pursing understanding grew into a large group of international researchers studying the area in methodicly detail, looking for solutions to practical problems. While it would be extreme to draw too close a simlarity between the orrigins of nuclear physics and network measurement there are some similarities. Network behaviour is difficult to predict and has, over the last 15 years or so, made an interesting field of study for a fairly small group of scientests. A body of methodoligies and understading has been built, primarily for no greater purpose than understanding network behaviour. There are some long term goals in mind but the problem has been too big for these to be too firmly fixed in the minds of the researchers. The basic ground work has been done and we believe that we stand at a potential explosion point for network measurement research, much like that of nuclear physics just before WWII. The most probably event that will stimulate the transition from interesting research to essential disipline is a catastrophic failure of the Internet. Like the explosion of "little Boy", such an event would, through success and tradegidy create an awareness of the importantance of network measurement that it does not currently enjoy.

Whether such an event will occur is unclear, that it is possible is certian. That it can be avoided by good quality network measurement is also equally without doubt. The future then, of network measuremnt research, depends on our success now. If we measure well, the science will grow slowly over years into a mature disipline. If we fail to warn of potential network failures, our failure will create an envionment where future success is essentail.

This rest of the paper explores the demands of network measurement through proposing some guidelines for quality measurement and examples of the work of two well know network measurement groups. The paper is organ-

Proceedings of the 2002 Symposium on Applications and the Internet (SAINT'02w)
0-7695-1450-2/02 $17.00 © 2002 IEEE

ised as follows: section 2 discusses the state of the art of network measurement and makes suggestions for moving beyound the current deployment model to a focus on greater accuracy of measurement. The following two sections describe the main projects of the (US) National Laboratory for Applied Network Research (NLANR) and the University of Waikato (WAND) Network Research Group. More details of this work can be found in [14] and [9]

2 Guidelines for Measurement

There are three methods of collecting data on network behaviour. These are:

- **passive measurement**, where a probe that records network activity is inserted into the network. Most commonly the probe is attached to a link between network nodes and summarises and records information about the traffic flowing on that link.
- **active measurement**, where the behaviour of the network is studied by sending data through the network and observing the results, including the time taken to send the data.
- **control monitoring**, where network control information, such as routing or network management information, is captured and analysed.

The three approaches have different focuses and it is often the case that more than one approach is required to develop an understanding of some aspect of a network's behaviour. Passive measurement observes the behaviour of a network at a specific point; it does not add to, or modify, the data carried by the network. Consequently, it has no impact on the behaviour of the network. A very detailed understanding of the behaviour at the point of measurement can be developed but it is difficult to gain an understanding of the network as a whole, or the end-to-end behaviour of the network. Passive measurement is often used to measure traffic characteristics, such as the mix of traffic by type or destination.

Unlke passive measurement active measurement involves sending traffic into the network. As a consequence, it lends itself to measuring parameters that reflect the service the network is offering to its users, including end-to-end parameters like round trip time and packet loss. However, the traffic inserted into the network may alter the behaviour of the network. This is particularly true for those parameters, such as available throughput, that are difficult to measure without sending large volumes of traffic.

Monitoring the control information of the network provides a ready source of information about those aspects of the networks behaviour which are described by data transfered as part of normal network operation. Parameters like link utilisation or route stability may be collected this way. Validation is difficult with control flow monitoring. Even assuming that the researcher can gain access to these flows (which can sometimes be administrately difficult) and that they include the information of interest, it may still be difficult to verify the accuracy of the information because its collection and transfer is outside the control of the researcher.

Many measurement activities are based arround gathering traces of network traffic, storeing them to disk and then archiveing and processing them. In the rest of this section we review this process from the perspective of quality. Although the focus here is not on real time processing and displaying, the advice is also applicable to that problem.

At each step of the network measurement process there are recurring themes that need to be addressed. The first of these is capacity. The bandwidth of modern networks and the need to capture data over long times, imposes stringent demands at every level on both bandwidth and storage capacity. This can be ameliorated by careful use of specialised hardware at critical points and by filtering data so that only what is essential is handed to the next stage. As network speeds continue to increase at rates greater than More's Law this problem will continue to become more severe because the balance of available computing resources to bits transfered worsens. This increases the need for the measurement community to move from naieve data collection, where the goal is unknown a prior, to deliberate measurement designed to capture the data actually needed for a know analysis. This is counter balanced by the need to maintain an archive of older data for historical analysises that are not currently predictable.

The next theme is confidence in the results, that they reflect actual network behaviour. An example of the type of problems that can arise is loss of data. This can occur at all stages of the measurement process from the point that data is captured on the wire to transfer to disk. Of more importance even than ensuring that that loss does not occur is having independent checks in place so that if loss does occur it is detected. It is interesting to note the the needs of network measurement vary from those of other areas of computer networks. A little loss does not matter very much in normal network operation but in measurement there is a danger that its effect will become amplified in later analysis steps. For example loss might occur during a burst in traffic, but it is precisely such moments that are important in assessing overflow statistics in queues. This has the interesting corollary that measurement equipment needs to be more reliable than standard network interface equipment.

Another important aspect of confidence is the need to maintain an audit trail of how data was captured and what has been done to it. Much of the data that is captured is archived and may be reprocessed long after initial capture.

Interpretation at this later date requires that careful and detailed information be kept of where and how the data was captured. Also, because there is always the possibility for anomalous and incorrect behaviours in the capturing software and hardware, it is necessary to be meticulous about recording which versions of software and hardware were used to capture the data as well as any options that were set during the capturing process. Any later processing of the data also needs to be recorded. For example, sometimes timestamps recorded on captured data are corrected after the fact. It is essential the software used to do the correction is recorded (and that copies of the software be kept as well).

The final theme that we will consider is access and security. The presence of large amounts of recorded traffic on disk is a juicy target for those who would like to snoop on networks - we have done the hardest part of hacking, that is gaining access to the network. Security is normally dealt with by: omitting sensitive data (for example recording only headers not full packet data); by encrypting parts of the data (for example the IP addresses); and by maintaining secure access to archived data.

2.1 Wire Level

The networks that are most interesting to analyse are those that carry large amounts of aggregated traffic. These links are also the most important to analyse from the point of view of telecommunications and service suppliers because they are costly and central to their operations. This means that measurement is being done on critical physical interfaces. Thus, even the act of tapping into the wire or fibre can be come a serious operation. This requires a good relationship with the owner of the physical interface.

There are typically two ways of tapping into high-speed links. One is to use a switch or router to duplicate the traffic to an output port that is used solely for measurement. The other is to tap into the physical wire directly. Diverting traffic has the disadvantage that in some cases it distorts traffic by increasing the demands on internal communication channels, so that it may not accurately reflect what is on any actual physical link. It has the advantage that it can be done without disrupting the traffic. The main disadvantage of physically tapping into a link is that it requires disrupting the link briefly, and that it may alter signal levels on the link. The main advantage is that it is guaranteed to be measuring what is on an actual link.

As described below GPS is often used to provide precise timing signals on for measurement hardware. This usually requires that there be an antenna installed on a roof and that a cable be taken from the GPS antenna to the machine room. This turns out to be remarkably difficult. In some cases, for security reasons, it is impossible. In others it is just expensive and awkward. This issue has become so difficult in practice that many groups have investigated alternatives to GPS including: using more accurate oscillators (including temperature controlled oscillators); synchronising with SONET clocks; and using timing signals from cell phone base stations. The latter, particularly from CDMA networks looks particularly prommising at this time.

2.2 Measurement Hardware

Standard NIC are normally the cheapest devices that provide interface to a given physical protocol. Measurements are sometimes made made with standard network interface cards (NICs), sometimes with some reprogramming of the NIC. However, this is not an ideal approach. The principal problem is that NICs are not designed for the accurate timestamping of cell or packet arrivals. A typical goal is to time IP packets with an accuracy and resolution that is approximately one tenth of a minimum sized packet, that is better than 275 nano-seconds at OC3, and 70 ns at OC12. With commercial NICs it is often not possible to get to within two orders of magnitude of this accuracy. Another problem is that no single manufacturer produceds a range of NICs that would gives consistent results for different network speeds and protocols.

2.3 Analysis

Analysis of traffic can proceed at a number of levels. For example, at the level of aggregate traffic it is possible to do studies of queuing behaviour. This is often done by fitting statistical models to the data. A different approach is to separate out different types of traffic and generate separate models for each type of traffic. See, for example, the case study separating out voice over IP traffic (see section 5. At the TCP level it is possible to extract separate measurements for the different packet types that occur in a session.

The final step in analysis is often to do simulations of network traffic. For example the voice over IP traffic seperated out of the traffic mix may become the input to a simulator evaluating different mixes of traffic than those that actually occured on the network. For an example of a simulation study by the WAND group see [15].

3 NLANR

The National Laboratory for Applied Network Research (NLANR) is a distributed research and support organisation focused on the high performance connection (HPC) community in the United States. This community is served by two National Science Foundation (NSF) approved high performance research networks. These are the vBNS[1] and Abilene[2] networks.

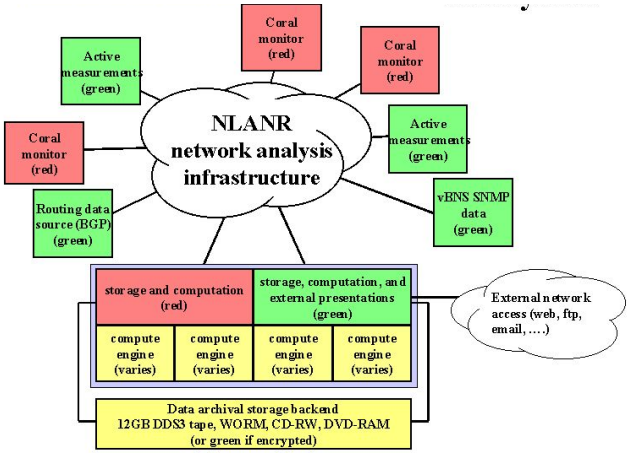

Figure 1. NAI Architecture

The Measurement and Operations Analysis Team within NLANR is developing a Network Analysis Infrastructure (NAI). It is intended that this infrastructure will provide both engineering and research support for the HPC community. Specifically, the goal of the NAI project is to create an infrastructure that will support measurements and analysis through the collection and publication of raw data, visualisation and analysis of network measurement. Currently the main focus is on:

- passive collection of header traces
- active measurement
- SNMP derived data
- BGP router based data
- presenting the results of analysis to the HPC community

To these ends, there are two well established projects and a number projects that are in the early stages of development.

3.1 NAI

The NLANR NAI infrastructure supports all three types of network monitoring (active, passive and control flow monitoring) as shown in figure 1. Newly collected data is stored on the machine `nai.nlanr.net` where local analysis and encoding of the data occurs. The data on this machine includes sensitive data like IP addresses. Desensitised data is copied to `moat.nlanr.net` for WWW publication. Researchers who need access to large quantities of data may be granted use of one or more of the compute engines in the infrastructure.

3.2 PMA

The NLANR passive measurement project utilizes OCXmon monitors. An OCXmon monitor is a rack mountable PC running the FreeBSD or Linux operating system. In addition to the PC components two measurement cards are installed in the PC and an optical splitter is used to connect the monitor to an OC3 (155mbps), OC12 (622mbps) or OC48 (1.5gbps) optical link. OC192 is under development and expected in the first quater of 2002. There are five types of measurement card available for OCXmon. These are:

- Fore ATM OC3 cards
- Applied Telecom ATM OC12
- DAG ATM/POS OC3/OC12/OC48 cards (developed by the University of Waikato WAND group [3]).
- Any FreeBSD supported FDDI card (we have used DEC DEFPA cards)
- Any 10/100 Ethernet card supported by FreeBSD.

OCXmon machines have two measurement cards installed in each monitor so that they can capture traffic in both directions of a full duplex connection. The total cost for an OC3mon based on the Fore measurement cards is under US$5000. Further details of the hardware are available at the project web page [4]. An Applied Telecom based OC12mon costs in the order of US$18,000. DAG cards are not currently commercially available but it is expected that an OC3/12 monitor based on DAG cards will cost around US$9,000. Endace Measurement Systems[5] are exppect to be maeketing the cards shortly.

Control software (originally developed at MCI by Joel Apisdorf in the case of the Fore and Applied Telecom cards) is downloaded to the measurement card to control the capturing of traffic. Using different versions of this software the ATM based monitors can be configured to capture different parts of the traffic they see. When an IP packet is sent over an ATM connection it is broken into 48 byte pieces to be sent in cells. Assuming the most common approach to breaking the packet into cells is being used (AAL5) the last cell of the packet is marked to indicate the packet is complete. The next cell (on the same virtual circuit) will be the first cell of the next packet. If there are no optional headers in use the 48 bytes of data in the first cell include the IP and TCP or UDP header but very little, if any, user data.

Currently, every three hours 90 seconds worth of header trace data is collected at each monitor. The start time within the hour is pseudo-random to avoid the risk of distortion created by regular network events. All monitors start at the same pseudo-random time so that events that are detected

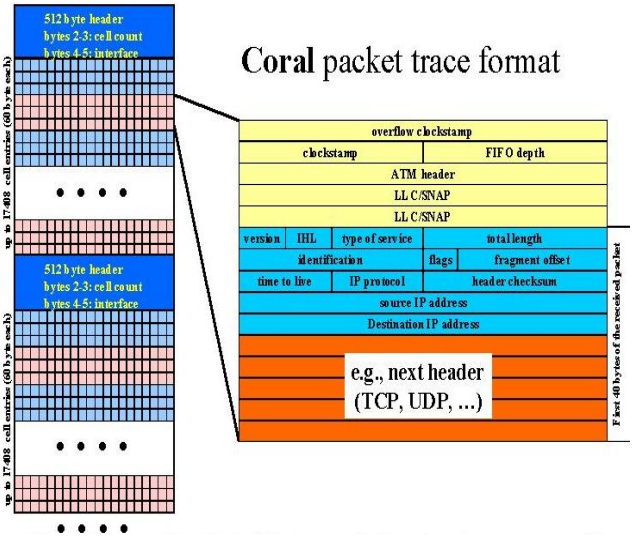

Figure 2. Trace Format

at different monitors may be synchronised. Following the trace collection several network analysis routines are run, and the results posted in the data cube[6].

A day's traces from a single site vary in size from a few tens of megabytes to several gigabytes, depending on the speed of the link being monitored and level of traffic. The disk space available in the NAI system allows about a months traces to be kept on line. Older traces are retired to tape but can be made available if they are needed for specific research needs. It is interesting to note that, so far, there have been few requests for historical data while there are many hits per day on the current trace files. We surmise that this is indicative of the state of networking research. The high rate of change of the Internet means that there are more important questions about the behaviour of today's networks than network researchers can answer. Consequently there is little research capacity available to investigate trends in historical data.

The OC3mon trace file format is shown in figure 2. Most of the fields in the cell entry are direct copies of the data in the captured cell. The first three fields are exceptions. The first two fields together record the time at which the cell was captured. Because the time stamp is added by the measurement card it is close to the wire time, relative to the time the card was last reset. Shortly after the measurement is started both cards are reset to synchronise the timestamps between the cards. Because most cards have separate clocks there may be a small amount of drift between the clocks over the period of the measurement. The DAG cards, which use a somewhat different file format, have a facility that allows a synchronising signal between the cards if very accurate synchronisation is required between the cards. The other additional header field is used to check for overload conditions where the monitor is unable to keep up with the workload.

After the measurement period some analysis programs are run on the captured files and the file and the analysis results are transfered to the central NAI machines.

The analysis that is done on the traces includes:

- bidirectional transaction analysis

  Statistics are generated on transactions where a transaction is a bidirectional sequence of packets with the same IP source, IP destination, IP protocol and port numbers (if appropriate). Transactions terminate when no matching packets are sent for 8 seconds.

  The results include statistics on individual transactions, aggregated across all transactions and a summary of the transactions with the highest throughput.

- flow analysis

  Similar statistics are also generated on flows where a flow is one (unidirectional) half of a transaction.

More details and example output from the analysis can be found on the WWW[6]. Other analysis programs are available through CAIDA's [7] CoralReef software. These include flows analysis, packet size and frequency histograms, packet size run lengths, protocol and port break down, host and autonomous system matrices, type of service breakdown and ASCII dumping of packets.

There are two foci for work in the immedate future. The first is a 'router clamp'. That is, we will instrument every connection into, and out of a router. This will allow a detailed study of the way traffic flows through a router, including agrigation studies. The second is dense instrimentation along a path, where every link on a path is instrimented. This will allow larger scale effects to be studied as a packet flows through the network from source to destination, including the sffects of access links and major agrigation points.

3.3 AMP

AMP[12] is NLANR's active measurement project. The focus is on making site to site measurements of round trip time (RTT), packet loss, topology and throughput across the National Science Foundation (NSF) approved HPC networks. At the time of writing around 130 monitors are deployed at NSF HPC awardee sites. This number is currently increasing by a few each month.

Each of these monitors sends a single ICMP packet to each of the others every minute and records the time to (or absence of) the reply. In addition, every 10 minutes the route to each other monitor is recorded using traceroute. Throughput tests can also be run between any pair

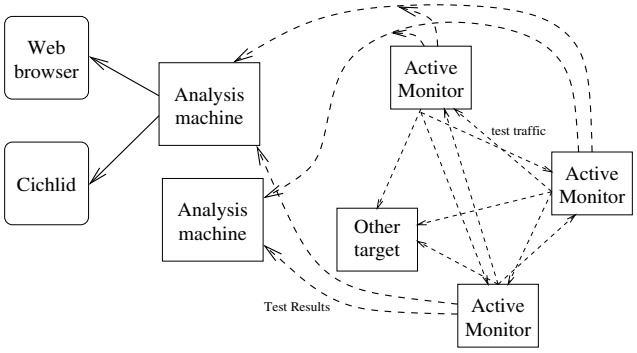

Figure 3. AMP Architecture

of the monitors using a web based throughput test request. (Throughput tests are only run on demand because of the high cost, in terms of network traffic, of running these tests.) The following throughput tests are available:

- Bulk TCP data transfer
- Bulk UDP data transfer
- ping -F
- treno
- iperf

The data collected is sent to the central AMP site, at the San Diego Supercomputer Center, for processing and publication via the WWW. To improve robustness, two central machines are used. The data is sent to each central machine independently by each monitor. This arrangement is shown in figure 3.

Because AMP measurements are continuous, there is no natural end time at which to send the collected data from the measurement machines to the central servers. In addition we want the system to have a near to real time nature with the data in the web pages for today being current to within a few minutes. This is important if the system is to be used as a diagnosis tool. To achieve these ends we have developed an active mirror system. This operates much like the daily mirror used on many FTP sites except that file changes are reflected on the mirror site more quickly. When a monitor is started it opens a TCP connection to each of the central machines. It then watches the last-modified date on the files in its directory tree and when a file is updated the changes are sent to each of the central machines. The process is fault tolerant so that if a central machine or a monitor fails, when it recovers all machines will be brought up to date. In addition to keeping the central sites current, this approach avoids a peaky transfer load that could overwhelm the central servers or disturb the measurements being taken by the monitor.

The data on the central web servers can be accessed from a web page that lists the monitor sites as hyperlinks. When a link is selected a table of the RTT and loss from that site to all the other sites is supplied. Again the site names are hyperlinks. If a site from the table is selected, RTT and loss data for that pair of sites is displayed as a year-to-date graph and a set of weekly graphs for all weeks this year. Further hyperlinks allow selection of a detailed display of any day, including the RTT by time of day and as a frequency distribution. The route data can be displayed in a tabular form (like the output of traceroute) or as a graphical plot using the Otter tool from CAIDA[7]. These displays are best understood by visiting the AMP web site, which is linked off the NLANR home page[8].

There are a large number of AMP monitors and consequently a very large number of pairs of monitors. Data is collected on the path between every pair of monitors and there are web pages for each pair. As described in section 3.1 this creates problems for people looking for interesting events in the data. We are addressing this through automated event detection[13].

4 Case Studies from WAND

4.1 Measurement Hardware

Our initial experiments in ATM data capture were made with re-programmed network interface cards (NICs). We discoverd that this was not an ideal solution, as explained in the section 2.2. So, we took the decision to design and build our own hardware. Our design aims in the Dag series were: a range of data capture boards with a consistent architecture, capable of handling data rates from 10 Mbps Ethernet to ATM and POS at OC48 and beyond; on board intelligence to filter the data before it is passed to the host processor; industry-standard interfaces - PCI and Compact PCI; a programmable and re-configurable structure to give the greatest flexibility to each design; highly accurate timing of cell or packet arrivals, referenced to a universal time standard; low cost.

The general architecture of the Dag series is shown in figure 4. In the lower speed Dags we have used ARM7 processors, the higher speed boards, Dag 3 and up, use a 233 MHz StrongARM. However even a 233 MHz processor cannot execute many instructions in the approximately 170 ns cell time of an OC48 network. Each board has the ability to receive periodic timing pulses. When used with a GPS antenna these pulses enable data to be time-stamped to an accuracy of $\pm 250ns$ to UTC. This is sufficient for us to time delays through ATM switches and to see timing jitter caused by SONET framing. In some cases the timing pulses are used only for local synchronisation. One board is designated the master and sends pulses to the other boards. One

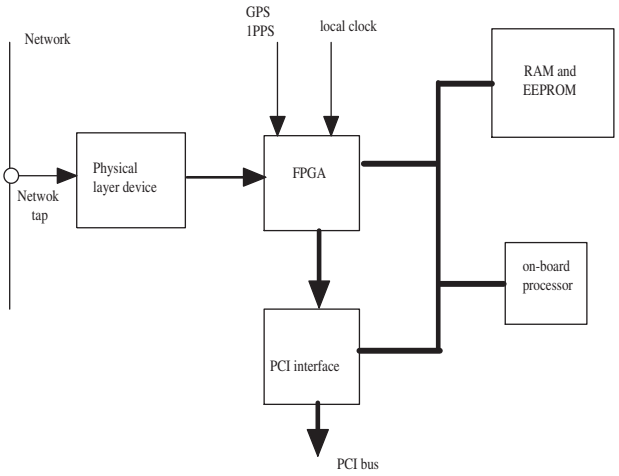

Figure 4. DAG Hardware Architecture

important use of this is to synchronise boards measuring the two directions of bidirectional ATM circuits. The relative timing between the boards is then accurate although they drift by approximately $\pm 0.5ppm$ to UTC which means a drift of about 1ms in a 15 minute run. One typical example of data filtering is where we need only to record packet headers for AAL5 encapsulated IP packets, on an ATM link. In almost all cases the entire header is contained with the first cell of each AAL5 PDU. This first cell is not marked in any way, but the last cell in each AAL5 PDU is marked by having a particular bit set in the payload type field of the cell header. AT OC3 and OC12 the StrongARM processor is quite fast enough to maintain state for each active virtual circuit, and to select only the first cell in each PDU for transfer to the PC. AT OC48 the processor is not fast enough, and has to be assisted with table lookup firmware within the FPGA. The board is programmable in two ways - in the code that runs on the on-board microprocessor, and in the configuration of the FPGA. This allows a single board to serve very different functions. For example, the Dag3.2 was initially intended to receive and analyse data on OC3 and OC12 ATM links. A simple reprogramming of the FPGA allows the board to measure both ATM and POS traffic. Recently, we have also programmed the board to act as a data source for network testing.

4.2 Decomposing WWW Response Times

Delays fetching a web page come from a number of different sources that can not all be attributed to simple round trip times (RTT). A question that is often asked is "Is the web slow because of the network or the servers?" The answer to this is important not only from a user point of view but also to know how improve the performance of the web. There is, for example, no point in adding networks capacity if servers are the main component of WWW delay. This analysis is an attempt to answer that question. Delays in fetching a web page come from four areas: server and stack processing delay; congestion and queueing delay; retries after message loss; physical network delay.

Estimating these components of delay individually relies on the observation that an HTTP transfer involves a number of different steps, which occupy the server CPU in varying amounts.

If a trace with accurate timestamps and many HTTP requests is collected it is possible to identify the separate components of WWW page fetching delay.[11]. Physical network delay (which includes speed of light delays as well as router and switching delays) can be estimated by looking at the minimum delay time for packets such as the ACK/SYN pairs. These pairs involve little CPU processing on the server. The physical network delay can then be subtracted from the other delay times. The time attributable to retries after message losses can be separated out because all packets both sent and received can be seen. Consequently any packet losses can be detected and separately accounted for. Finally, the effects of server CPU delays can be separated out by using the fact that some messages involve server CPU and disk activity and others do not. This allows the delays for the different classes to be counted separately and allows queueing and congestion to be estimated as well as server delay. Although this work is still in progress early results, shown in figure 5, indicate that for the NZIX the delays experienced by most HTTP requests are dominated by server delays. In only 3 of the 40 connections analysed is the queueing delay greater than the server delay. This data strongly suggests that the bulk of experienced network delays are caused by delay on servers. This result is even more surprising because New Zealand is physically remote from many web servers and it might be expected that access to many sites would be dominated by physical delay and queueing delay from the intermediate hops.

5 Detection of Voice over IP

The use of the Internet to carry voice over IP (VoIP) communication has been increasing recently . There is concern that interaction between VoIP, which uses UDP, and other Internet traffic, most of which uses TCP, may cause congestion for TCP users or even congestion collapse. This problem can occur because TCP reduces its transmission rate when congestion is detected while UDP does not.

The WAND group's VoIP studies[10] focus on H.323, the ITU standard for VoIP. Multiple vendors already use this protocol, and many other VoIP systems are being converted to use H.323. The goal is to create an accurate simulation

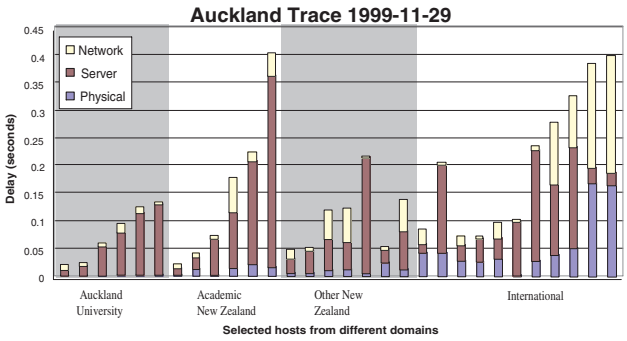

Figure 5. TCP Connection Delay Sources

model of a mixed TCP/UDP environment. To build such a model it is necessary to capture sample VoIP sessions and either create a synthetic data source with the same characteristics, or use the captured data to generate a trace that can drive the simulator. Two options are available for obtaining sample data for the simulation. The first is to measure sample conversations in a lab environment. The advantages of this are complete control of the systems and protocols used. The second option is to capture a set of IP header traces and, using the signature of H.323 sessions, to identify data that is likely to be voice traffic. This approach will give a more varied and realistic picture of VoIP traffic for a number of reasons. First there are multiple implementations of both H.323 compliant programs, and voice compression systems. Second, different users of VoIP applications have different connection and machine speeds. Each user will therefore use different data rates and compression techniques.

Identifying H.323 completely accurately without any IP payload and no prior knowledge is impossible. Without being able to view the contents of the connection setup and call control packets it is impossible to tell exactly what the two machines are doing, and what the data passing between them actually means. However H.323 has a distinctive enough signature to make a heuristic approach possible. Currently a two-stage identification system is used. The first step looks for pairs of IP addresses that communicate on two known TCP ports: 1503; and 1720. These ports are the listen ports specified by H.323. The next step relies on recognising the use of the Real Time Protocol (RTP), RFC1889, whose use is also specified by H.323. This protocol is responsible for the transfer of the voice data itself. H.323 sets up dynamic UDP ports for RTP to run over, but RTP specifies the use of two adjacent UDP ports. The first port will be an even number, and will contain the data connection. This port will therefore have a large number of regular (often fixed) sized packets. The next (odd numbered) port will contain a small number of packets used for call control information. If two IP addresses are communicating in these ways it is likely, although not certain, that the data represents an H.323 session.

6 Conclusion

The science of network measurement has matured over recent years, but its contribution to the Internet is still to fully be recognised. Much good work has been achieved but we believe that network measurement researches need to move into a new phase of measurement, with a stronger focus on answering real problems and making practical improvements to the network. “The fields are ripe to the harvest, but the workers few” –Luke 10:2.

References

[1] http://www.vbns.net/.

[2] http://www.internet2.edu/abilene/.

[3] http://atm.cs.waikato.ac.nz/wand/.

[4] http://moat.nlanr.net/.

[5] http://www.endace.co.nz/.

[6] http://moat.nlanr.net/Datacube.

[7] http://www.caida.org/.

[8] http://www.nlanr.net/.

[9] J. Cleary, I. Graham, A. McGregor, M. Pearson, I. Ziedins, J. Curtis, S. Donnelly, J. Martens, and S. Martin. High precision traffic measurement by the wand research group. *IEEE Communication Magazine.*, 2001. *accepted*.

[10] J. Curtis, J. Cleary, A. McGregor, and M. Pearson. Measurement of voice over ip traffic. In *PAM2000 Passive and Active Measurement Workshop*, pages 43–59, Apr. 2000.

[11] H. Martin, A. McGregor, and J. Cleary. Analysis of internet delay times. In *PAM2000 Passive and Active Measurement Workshop*, pages 141–148, Apr. 2000.

[12] A. McGregor and H.-W. Braun. Balancing cost and utility in active monitoring: The amp example. In *The Global Internet Summit –Inet2000, Yokohama, Japan*, July 2000.

[13] A. McGregor and H.-W. Braun. Automated event detection for active measurement systems. In *PAM2001 A Passive and Active Measurement Workshop, Amsterdam, The Netherlands*, pages 23–32, May 2001.

[14] A. McGregor, H.-W. Braun, and J. Brown. The NLANR NAI network analysis infrastructure. *IEEE Communication Magazine: special issue on network measurement*, pages 122–128, May 2000.

[15] M. Pearson and A. McGregor. *A simulation study of network architectures to support HTTP Traffic on Symetric high-bandwidth*delay circuits*, chapter 3, pages 19–25. C.S.R.E.A. Press, 2000.